In vitro kinetic study of the squalestatin tetraketide synthase dehydratase reveals the stereochemical course of a fungal highly reducing polyketide synthase†

Emma Liddle,[a] Alan Scott,[a] Li-Chen Han,[a] David Ivison,[a] Thomas J. Simpson,[a] Christine L. Willis[a] and Russell J. Cox*[abc]

Six potential diketide substrates for the squalestatin tetraketide synthase (SQTKS) dehydratase (DH) domain were synthesised as N-acetyl cysteamine thiolesters (SNAC) and tested in kinetic assays as substrates with an isolated DH domain. 3R-3-hydroxybutyryl SNAC 3R-16 was turned over by the enzyme, but its enantiomer was not. Of the four 2-methyl substrates only 2R,3R-2-methyl-3-hydroxybutyryl SNAC 2R,3R-8 was a substrate. Combined with stereochemical information from the isolated SQTKS enoyl reductase (ER) domain, our results provide a near complete stereochemical description of the first cycle of beta-modification reactions of a fungal highly reducing polyketide synthase (HR-PKS). The results emphasise the close relationship between fungal HR-PKS and vertebrate fatty acid synthases (vFAS).

Iterative fungal polyketide synthases (PKS) are responsible for the biosynthesis of complex and often biologically active natural products such as squalestatin S1 1 a potent inhibitor of squalene synthase,[1,2] and lovastatin 2,[3] an inhibitor of human HMG-CoA reductase. These PKS are Type I systems in which several individual catalytic domains are covalently linked to form a mega-complex of ca 4200 KDa.[4,5] Understanding the selectivity and programming of these systems is important because repro-gramming them could lead to the systematic creation of new bioactive materials. In order to achieve this an understanding of the individual catalytic domains and their intrinsic selectivities is required.

The C_{10} side-chain of 1, known as squalestatin tetraketide 3, is synthesised by a highly reducing (HR) iterative PKS called squalestatin tetraketide synthase (SQTKS).[6,7] It consists of an acyl carrier protein (ACP) which holds the growing polyketide chain, a b-ketoacyl ACP synthase (KS) which catalyses a Claisen condensation between malonyl ACP and acyl-KS and an acyl transferase (AT) which loads acetyl starter and malonyl extender units from CoA onto the PKS. In addition, SQTKS contains: b-ketoacyl ACP ketoreductase (KR); b-hydroxy acyl ACP dehydratase (DH); and enoyl ACP reductase (ER) domains which process the b-carbonyl formed by the KS. Finally, a C-methyl transferase (C-MeT) domain is responsible for adding a methyl group derived from S-adenosyl methionine. SQTKS thus contains a full set of active HR-PKS domains.

SQTKS shows sequence homology to vertebrate FAS (vFAS, Fig. 1).[8] This similarity even extends to the position of a C-MeT domain which is inactive in vFAS, but which acts during the first and second rounds of chain processing by SQTKS.[9] vFAS produces fully saturated linear 16–18 carbon chains, whereas SQTKS produces the dimethylated and unsaturated 8-carbon chain 3. SQTKS thus displays a complex programme in which the activities of the individual catalytic domains can be varied (Scheme 1). Our approach to study the programming mechanisms of fungal HR-PKS is to examine intrinsic selectivities of isolated catalytic domains. For example, we recently reported on the

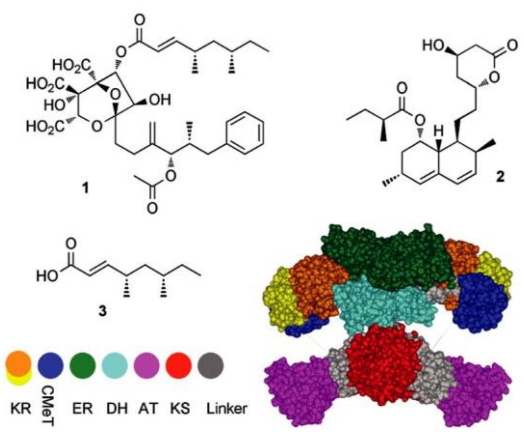
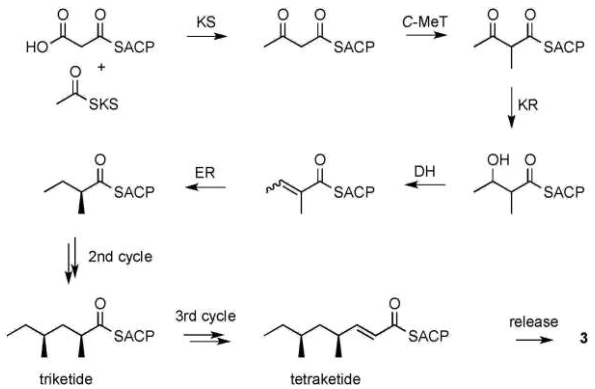


Fig. 1 Structures of fungal highly reduced polyketides, and the structure of dimeric vFAS (PDB 2VZ8)[8] which is homologous the squalestatin tetraketide synthase (SQTKS). See text for abbreviations.

[a] School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK
[b] Institut fu¨r Organische Chemie, Leibniz Universita¨t Hannover, Schneiderberg 1b, 30167 Hannover, Germany. E-mail: russell.cox@oci.uni-hannover.de
[c] BMWZ, Leibniz Universita¨t Hannover, Schneiderberg 38, 30167, Hannover, Germany



Scheme 1 The chemical reactions catalysed by SQTKS.

chemo- and stereo-selectivity of the isolated SQTKS ER domain.[10] Here we describe work to extend this study to the isolated DH domain of SQTKS.

SQTKS is a megacomplex of 284.4 KDa encoded by the phpks1 gene. We have been unable to obtain it as a single soluble protein. However, by systematic variation of possible start and stop positions for PCR from an intron-free phpks1 template, we were able to create an open reading frame which reliably produces soluble SQTKS DH protein when expressed in E. coli BL21 with an N-terminal his_6 tag. The DH protein of the expected 38.0 kDa was purified to homogeneity by nickel affinity and gel-filtration chromatography. Calibrated gel filtration indicated that the DH exists primarily as a monomer (see ESI†). The isolated DH was unstable in unmodified buffers, precipitating rapidly even at low temperatures. However, rapid removal of imidazole used for the nickel ion chromatography and use of a buffer containing 10% glycerol, 50 mM Tris pH 8.0, 150 mM NaCl and 100 mM L-arginine and L-glutamic acid dramatically improved protein solubility and stability.

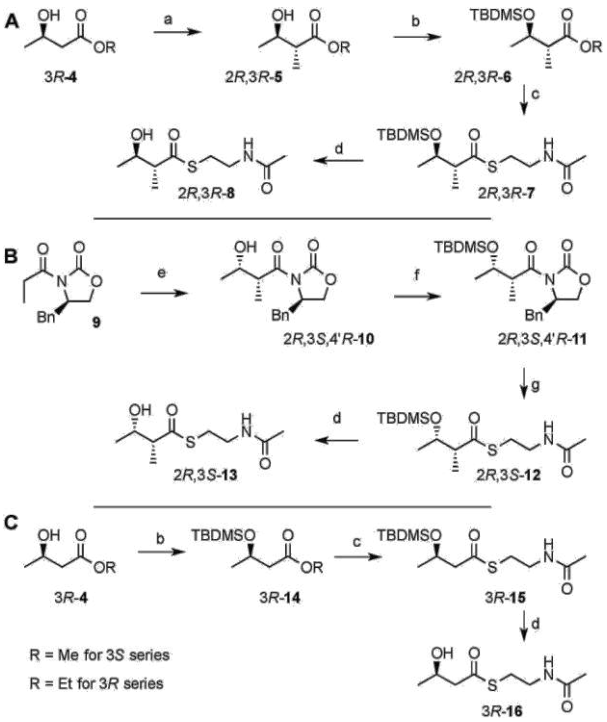
N-Acetyl cysteamine (NAC) is a truncated form of the phosphopantetheine (PP) cofactor which attaches acyl PKS intermediates to the ACP domain, and SNAC thiolesters are often used as PP surrogates for in vitro studies of PKS enzymes,[11,12] including DH domains.[13,14] We thus selected SNACs as targets for substrate synthesis (Scheme 2).

The anti diketide SNAC 2R,3R-8 was made by a route involving Fra´ter-Seebach methylation[15] of commercially available enantiopure 3-hydroxy butyrate 4 (Scheme 2A) to give 5. This was O-protected with TBDMS to give 6, which was in-turn hydrolysed to its corresponding acid and coupled to HSNAC to give the protected diketide 7. Acidic deprotection then yielded 2R,3R-8.

The syn diketide SNAC 2R,3S-13 was made using Evans asymmetric aldol chemistry[16] to give the known syn aldol product 2R,3S,4^{0}R-10 (Scheme 2B). This was again O-protected with TBDMS to give 11, which was hydrolysed and processed to the protected SNAC 2R,3S-12. Acidic deprotection then yielded 2R,3S-13.

The non-methylated diketide 3R-16 was made from 4 by a similar protection, thiolesterification and deprotection route (Scheme 2C). The enantiomers of all the diketides were made from enantiomeric starting materials using identical methods.

DH activity was assayed using LCMS (see ESI†) to measure substrate consumption and product formation. Assays were set



Scheme 2 Synthesis of potential SNAC substrates for the SQTKS DH domain: (a) LDA (2 eq.), 78 1C, then MeI; (b) TBDMSOTf, pyridine, 0 1C; (c) aq. LiOH, 60 1C, then HSNAC, EDCI, DMAP, 0 1C; (d) THF/H_2O/HOAc, RT, 5 days; (e) Bu_2BOTf, CH_2Cl_2, Et_3N, 78 1C, then CH_3CHO; (f) TBDMSCl, CH_2Cl_2, DMAP, imidazole, RT; (g) LiOOH, H_2O, RT, then HSNAC, EDCI, DMAP, 0 1C. TBDMS = ($Si(Me_2)CMe_3$).

up to include DH protein, substrate and buffer at 30 1C in 100 mL assay volume, and 20 mL aliquots were taken at time points and quenched in CH_3CN (60 mL). Protein was precipitated by centrifugation and the supernatant was examined directly by LCMS.

In order to maximise sensitivity, single ion monitoring was applied for substrate and product peaks and peak areas were integrated. The peak integrals were calibrated vs. known concentrations of substrate and product. Initial rates were deter-mined by plotting product concentration vs. assay times, and variation of initial substrate concentrations allowed the estima-tion of kinetic parameters (Fig. 2). The diketide 2R,3R-8 was dehydrated by the DH to give exclusively the E-olefin product tigloyl SNAC 17, but none of the other 2-methyl diketides showed any turn-over. Of the non-methylated diketides, only 3R-16 was a substrate, although much slower than 2R,3R-8. The ability of the non-substrates 2S,3S-8 and the enantiomers of 13 to act as inhibitors of the DH was investigated. However, addition of each of these compounds to assays containing the substrate 2R,3R-8 showed no appreciable decrease in rate when added in mM concentrations (see ESI†).

Despite having access to soluble protein we were unable to grow satisfactory crystals of the isolated DH domain. In lieu of other structural information we built a model of SQTKS DH based on the known crystal structures of DH proteins from other Type I systems reported in the literature. These form

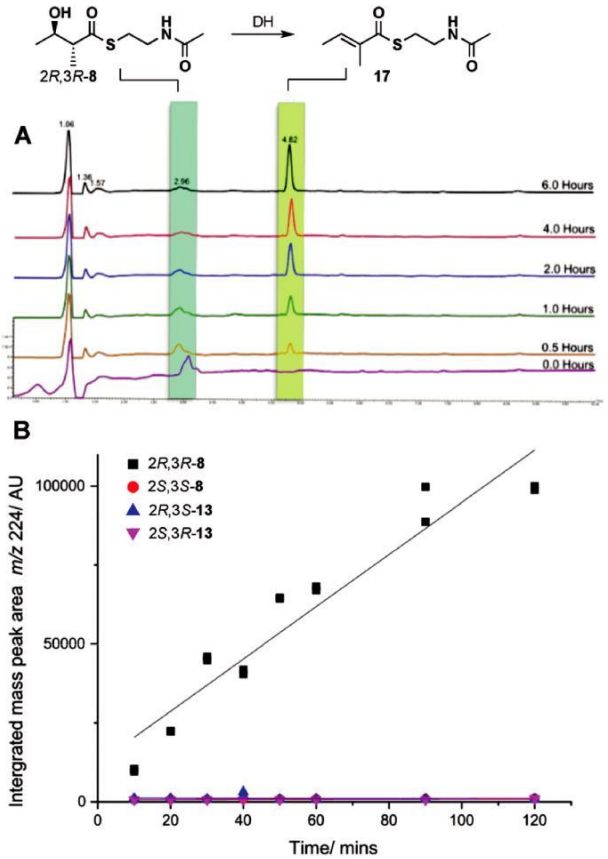


Fig. 2 Kinetic data for SQTKS DH reactions. (A) Raw time course data showing increase in tigloyl SNAC 17 concentration; (B) comparison of DH reaction rate for the 4 diketide stereoisomers.

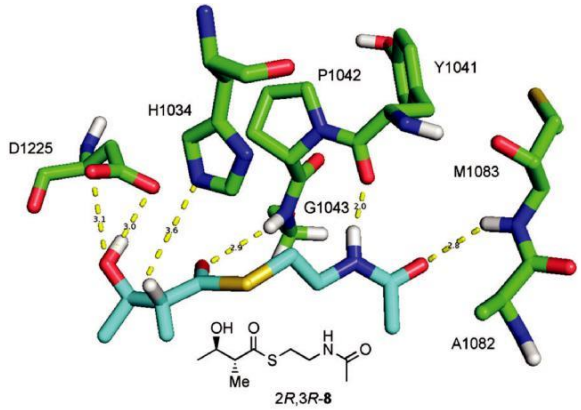


Fig. 3 Key modelled interactions between the substrate 2R,3R-7b (cyan) and the DH (green).

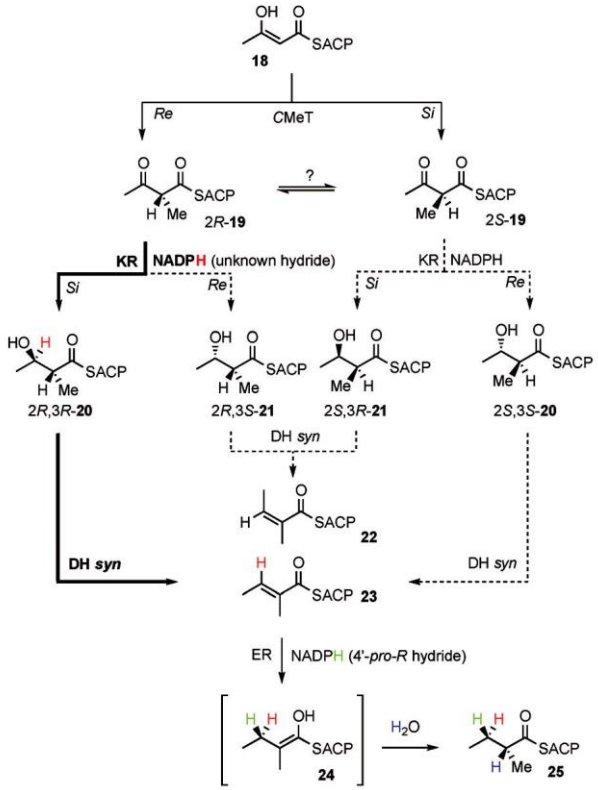
distinctive double hot-dog folds.[17] In particular the DH domain from CurF,[18] a modular Type I PKS, formed an appropriate template for the assembly of a model by the SwissModel threading server.[19] Comparison of the results showed that the backbone atoms of the SQTKS DH model and CurF-DH had only 1.4 Å root mean square deviation (RMSD). Almost all of the observed deviation was concentrated on the periphery of the model structure, and examination of the conserved active site aspartic acid (D1225) and histidine (H1034) residues showed that these amino acids are located in the same positions in the CurF DH and the model (see ESI†). In addition a highly conserved Y/FP motif (Y/F1041-P1042) is also preserved in the model.

The diketide substrate 2R,3R-8 was then docked into the DH active site using a combination of manual positioning (PyMol)[20] and optimisation and energy minimisation using the YASARA force field.[21] For 2R,3R-8, the docked model shows that the thiolester carbonyl oxygen is within hydrogen bonding distance of the backbone NH of G1043 (2.9 Å), and aligns with the dipole of the helix between G1043 and M1055 of the DH model. The SNAC NH is positioned within hydrogen bonding distance of the backbone carbonyl of conserved Y1041 (2.0 Å); while the SNAC carbonyl is 2.8 Å from backbone NH of M1083. These interactions locate the SNAC in the entrance to the reaction cavity and would presumably perform a similar role with the pantetheinyl-ACP-bound substrate in the native complex. The two methyl groups of 2R,3R-8 occupy open space and allow the a-proton and b-hydroxyl to approach the conserved histidine and aspartate residues. The closest distance between the a-proton and the nearest nitrogen of H1034 is 3.6 Å, while the distance between the b-hydroxyl and the oxygen of D1225 is 3.0 Å. The geometry of the interactions is consistent with a syn abstraction of water to give the observed E-2-methylbut-2-enoyl (tigloyl) SNAC product 17. Attempts to dock the other stereoisomers of the diketide did not give satisfactory results (Fig. 3).

Our experiments report the first in vitro studies of the stereoselectivity of an isolated DH domain from an iterative Type I PKS. Using a kinetic assay we measured the K_M (4.5 mM) and k_{cat} values (0.063 min ^{1}). While these values have little absolute meaning, they are comparable with values measured for other DH proteins. For example Aldrich, Smith and coworkers reported K_M values in the same range for a KR-DH didomain from module 2 of the pikromycin modular PKS (pikKR2-DH2) acting on triketide mimics, although their k_{cat} values are ten-fold higher.[13]

Only one 2-methylated diketide, 2R,3R-8, is accepted as a substrate for SQTKS DH, with no dehydration activity observed for its enantiomer 2S,3S-8 or either of the syn diastereomers 13. Since these stereoisomers show no measurable substrate or inhibition activity it seems unlikely that they can be bound at the DH active site, also supported by the failure to generate satisfactory docked models of these isomers. However the non-methylated diketide 3R-16 is a substrate.

In the active SQTKS the ACP-bound 2R,3R-diketide 20 is created by reduction of a 3-oxo diketide 19 by the KR domain using NADPH as the cofactor (Scheme 3). Our results strongly suggest that the SQTKS KR releases 3R substrates, and thus it must reduce the 3-oxo group of its substrate 19 by 3-Si addition of hydride. Since racemisation at the 2-position of the diketide is strongly disfavoured after reduction of the 3-oxo group, this observation also suggests that the KR accepts and releases 2R-methylated diketides (e.g. 2R-19, Scheme 3). However, because facile epimerisation of 2-methyl-3-oxo substrates such as 19 is likely, it is not yet possible to infer the stereochemical preference of the C-MeT without further experiments. Our previous results



Scheme 3 Stereochemical course of KR, DH and ER domains of SQTKS.

have shown that the SQTKS ER domain can process both Z 22 and E 23 substrates. However since the DH can only provide E-diketides it appears that the ER's ability to accept Z-olefins is merely adventitious.[10]

In our earlier study of the stereoselectivity of the ER domain we showed that the stereochemical preferences at the b-carbon are identical for SQTKS ER and vFAS ER, in terms of both the cofactor itself (transfer of 4^0-pro-R Hydrogen) and the substrate (addition of hydride to the 3-Re face).[10] The results of this study also show that the SQTKS DH has exactly the same stereochemical selectivity as the vFAS DH which dehydrates 2R,3R substrates 20 to give E-products 23 by syn elimination.[22] Even though the SQTKS substrate is methylated at the 2-position, the 2R stereochemistry ensures that the 2-pro-S proton is removed during reaction. Our model structure shows that the 2R,3R substrate aligns with the active site residues such that syn elimination gives the observed E-product. The active site residues involved, H1034, D1225, Y1041 and P1042 are conserved between the SQTKS and vFAS sequences.

Finally, the SQTKS KR domain also operates with the same stereochemical preference as the vFAS KR.[22] Although we have not yet been able to show which of the cofactor 4^0-hydrides is transferred by KR, the reduction does occur at the 3-Si face of the substrate. Thus our studies show that SQTKS shares more than just sequence homology and domain organisation with vFAS: its fundamental mechanisms for substrate reaction and stereoselectivity are also preserved and reinforce the idea that fungal hr-PKS and vFAS evolved from a common ancestor. Our current work focusses on determining the stereochemical pre-ference of the CMeT domain and attempts at engineering SQTKS to rationally change its selectivity.

We thank EPSRC (EP/F066104/1) for LCMS equipment, BBSRC (BB/I003355/1) for funding AS and BrisSynBio Synthetic Biology Research Centre, (BB/L01386X/1) for funding L-CH. E. L. and D. I. were funded by the School of Chemistry, University of Bristol. We thank Dr J. E. Nettleship at the Oxford Protein Produc-tion Facility (OPPF) for assistance with protein production. Hao Yao and Oliver Piech (LUH) are thanked for technical assistance.

Notes and references

1 B. Bonsch, V. Belt, C. Bartel, N. Duensing, M. Koziol, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, Chem. Commun., 2016, 52, 6777–6780.
2 P. J. Sidebottom, R. M. Highcock, S. J. Lane, P. A. Procopiou and N. S. Watson, J. Antibiot., 1992, 45, 648–658.
3 L. Hendrickson, C. R. Davis, C. Roach, D. K. Nguyen, T. Aldrich, P. C. McAda and C. D. Reeves, Chem. Biol., 1999, 6, 429–439.
4 R. J. Cox, Org. Biomol. Chem., 2007, 5, 2010–2026.
5 Y.-H. Chooi and Y. Tang, J. Org. Chem., 2012, 77, 9933–9953.
6 R. J. Cox, F. Glod, D. Hurley, C. M. Lazarus, T. P. Nicholson, B. A. M. Rudd, T. J. Simpson, B. Wilkinson and Y. Zhang, Chem. Commun., 2004, 2260–2261.
7 E. J. Skellam, D. Hurley, J. Davison, C. M. Lazarus, T. J. Simpson and R. J. Cox, Mol. BioSyst., 2010, 6, 680–682.
8 M. Leibundgut, T. Maier, S. Jenni and N. Ban, Curr. Opin. Struct. Biol., 2008, 18, 714–725.
9 K. M. Fisch, A. M. Bailey, W. Bakeer, A. A. Yakasai, Z. Song, J. Pedrick, Z. Wasil, C. M. Lazarus, T. J. Simpson and R. J. Cox, J. Am. Chem. Soc., 2011, 133, 16635–16641.
10 D. M. Roberts, C. Bartel, A. Scott, D. Ivison, T. J. Simpson and R. J. Cox, Chem. Sci., 2017, DOI: 10.1039/c6sc03496a.
11 R. C. Harris, A. L. Cutter, K. J. Weissman, U. Hanefeld, M. I. C. Timoney and J. Staunton, J. Chem. Res., 1998, 283; I. E. Holzbaur, A. Ranganathan, I. P. Thomas, D. J. Kearney, J. A. Reather, B. A. Rudd, J. Staunton and P. F. Leadlay, Chem. Biol., 2001, 8, 329–340.
12 D. E. Cane, F. Kudo, K. Kinoshita and C. Khosla, Chem. Biol., 2002, 9, 131–142.
13 Y. Li, G. J. Dodge, W. D. Fiers, R. A. Fecik, J. L. Smith and C. C. Aldrich, J. Am. Chem. Soc., 2015, 137, 7003–7006.
14 N. Kandziora, J. N. Andexer, S. J. Moss, B. Wilkinson, P. F. Leadlay and F. Hahn, Chem. Sci., 2014, 5, 3563–3567; G. Berkhan and F. Hahn, Angew. Chem., Int. Ed., 2014, 53, 14240–14244.
15 G. Fra´ter, U. Mu¨ller and W. Gu¨nther, Tetrahedron, 1984, 40, 1269–1277.
16 M. A. Walker and C. H. Heathcock, J. Org. Chem., 1991, 56, 5747–5750.
17 J. W. Labonte and C. A. Townsend, Chem. Rev., 2013, 113, 2182–2204.
18 D. L. Akey, J. R. Razelun, J. Tehranisa, D. H. Sherman, W. H. Gerwick and J. L. Smith, Structure, 2010, 18, 94–105.
19 M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli and T. Schwede, Nucleic Acids Res., 2014, 42, W252–W258.
20 The PyMOL Molecular Graphics System, Version 1.8 Schro¨dinger, LLC.
21 E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker and K. Karplus, Proteins, 2009, 77, 114–122.
22 V. E. Anderson and G. G. Hammes, Biochemistry, 1984, 23, 2088–2094.